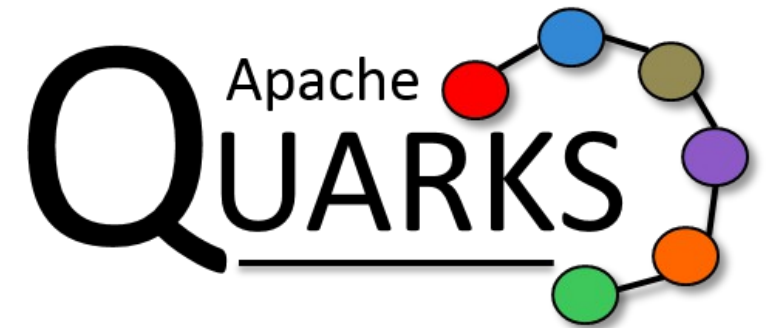# Apache Quarks for Developers

April 13, 2016

Apache Quarks is currently undergoing Incubation at the Apache Software Foundation.

Apache Quarks Community

- A **community** for accelerating Edge Analytics
  - Open Source, incubating at Apache Software Foundation
    - http://quarks.incubator.apache.org/
    - http://wiki.apache.org/incubator/QuarksProposal
- A programming SDK with functional flow API for streaming analytics
    - Initial support for Java 7,8 & Android,
    - Goal is to support multiple languages with priorities driven by the community
- A modular, lightweight and embeddable runtime

What is the Edge?

- Constrained compute platform
  - Due to cost, weight restrictions, space constraints, …

- Limited connectivty to central systems
  - Limited by expense, bandwidth
  - Periods of being disconnected

- Access to sensors for system being analyzed
  - Directly or through a bus
  - Potential to control system

- Can be mobile or static

- Expected to be thousands to millions of devices
  - Internet of Things

Edge Examples

- Vehicle
  - Car, truck, race car, train,bus, boat, drone, …
  - Analyze engine sensors to predict/reduce chance of failure
  - Mobile, may lose connectivity

- Building
  - HVAC, climate, energy use, motion sensors, …

- Server in machine room
  - Analzye load, cpu temps, rack temps

- Raspberry Pi with a couple of sensors
  - Cheap $5+

- Smartphone

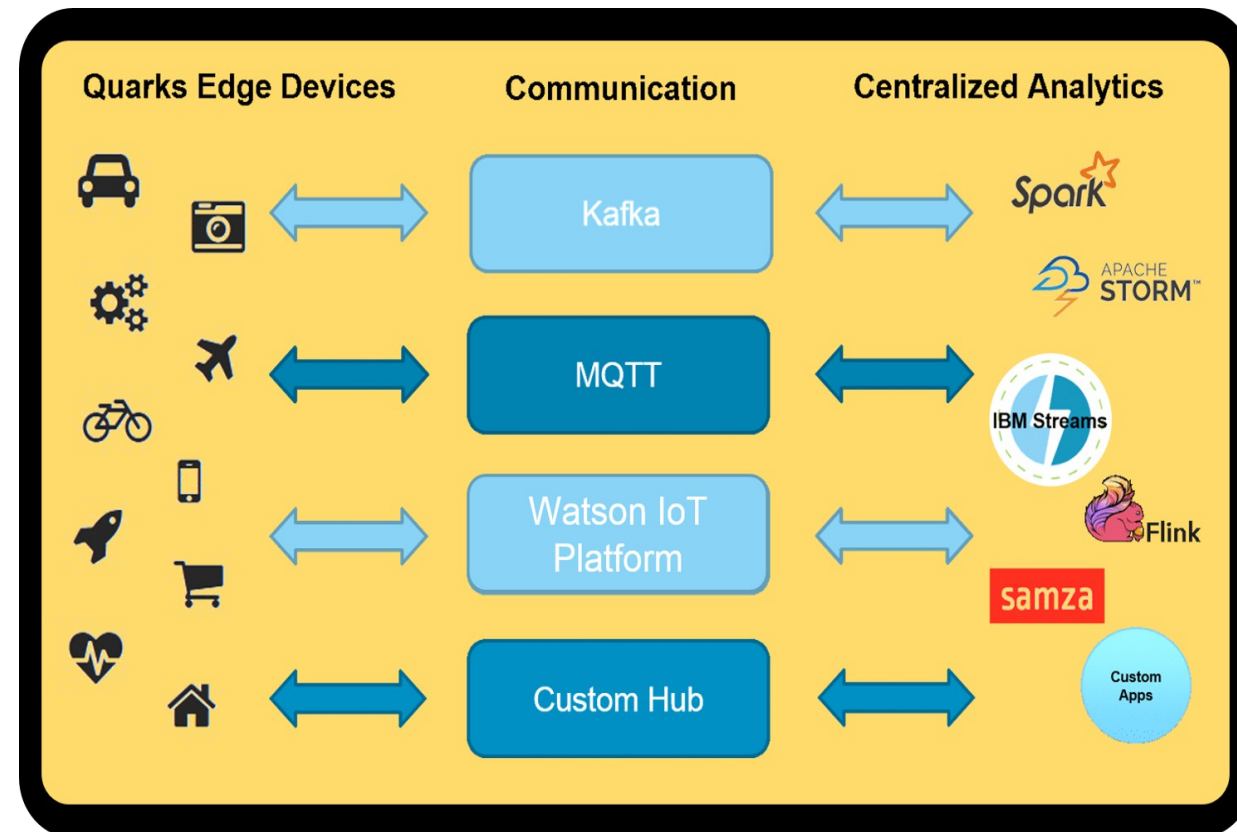# What is the value of analytics at the edge?

- Reduce communication cost
  - Send relevant data when an event of interest occurs
  - Heartbeats alone may not contain enough data or be too late to take action
- React locally to events
  - More intelligent decision making on the device
  - Execute analytics while disconnected
- Collaborate with related devices
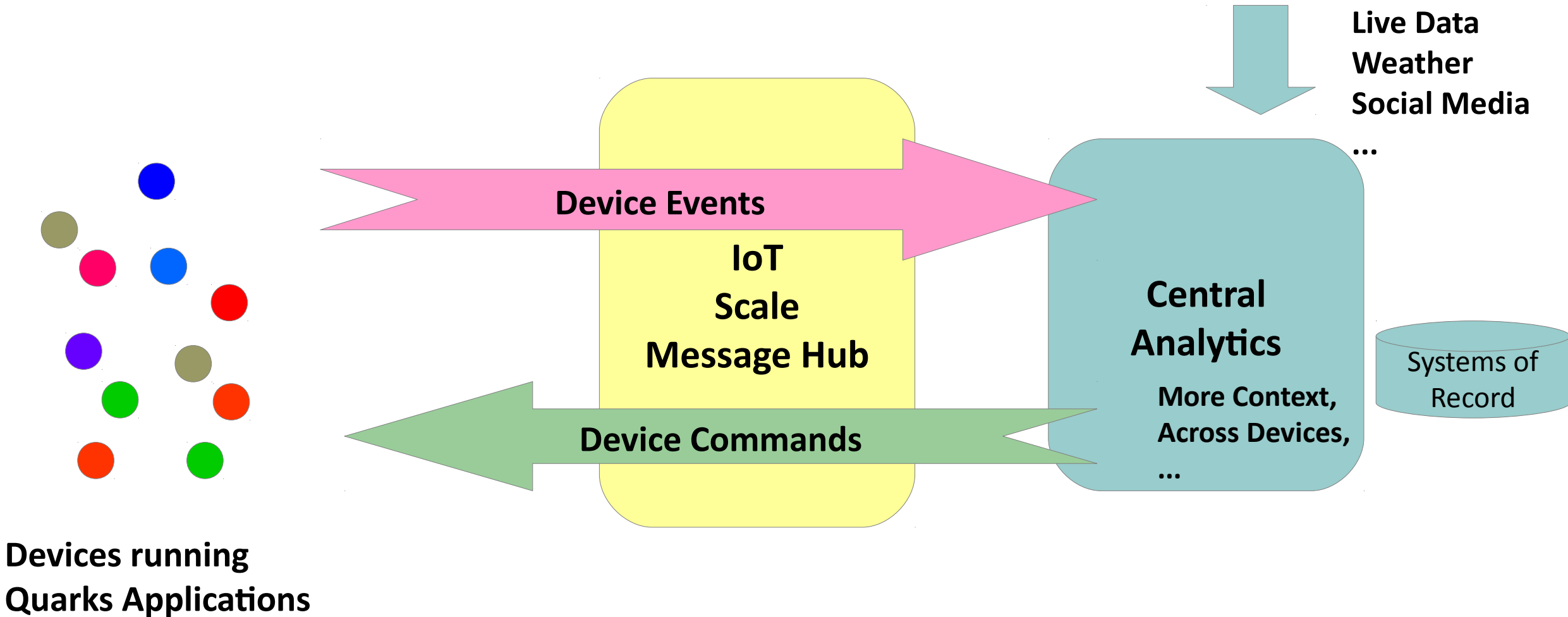 - Learn from devices with similar characteristics
  - or location
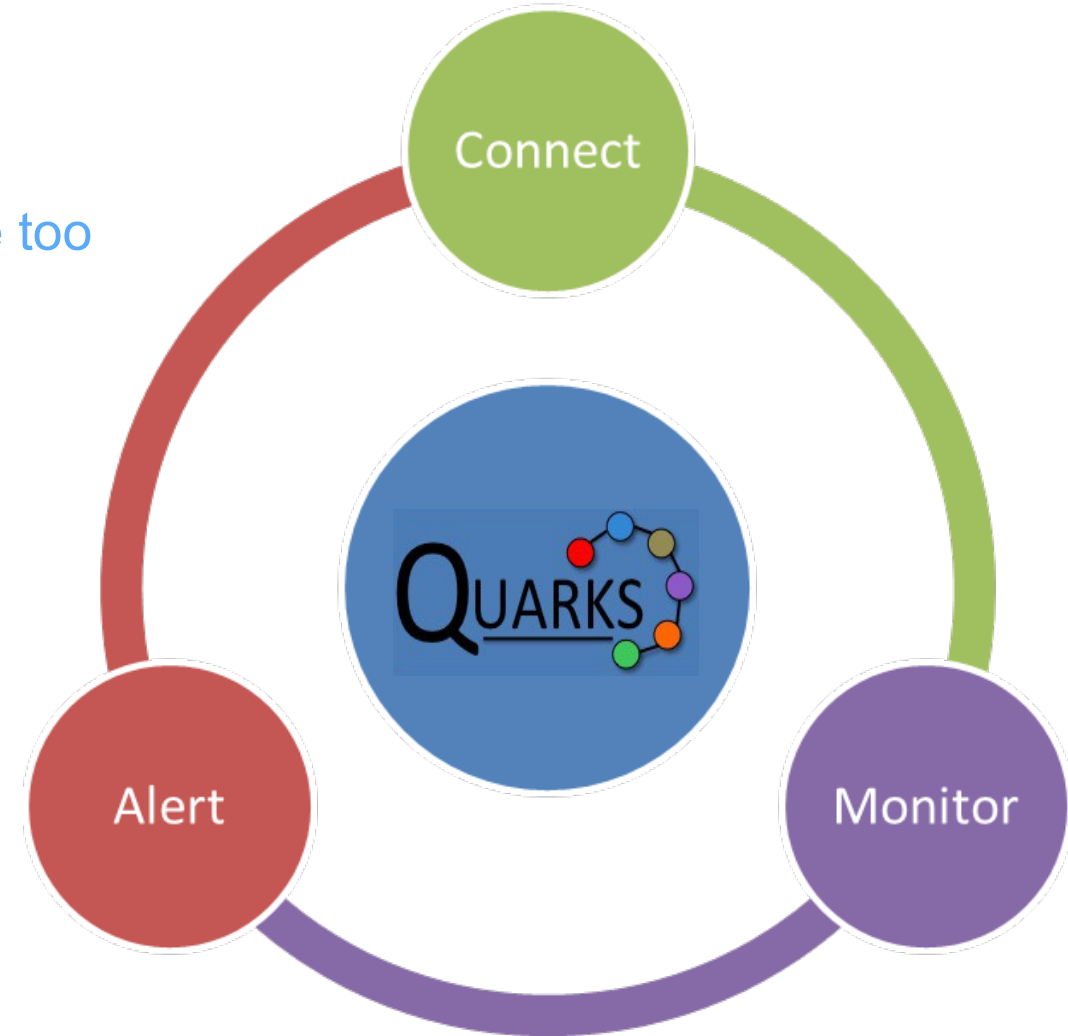
Integrated with Centralized Analytics

- Integrates with centralized analytics systems through IoT scale message hub
  - Any hub
  - Any central system
-

Control loop through Central Analytics

Device Events

**IoT
Scale
Message Hub**

Device Commands

**Live Data
Weather
Social Media
...**

**Central
Analytics**

**More Context,
Across Devices,
...**

Systems of
Record

**Devices running
Quarks Applications**

Device Model

- Quarks applications
  - Send *device events* to be centrally analyzed
    - engine data when coolant temperature is increasing
    - GPS location updates
    - Traction control activated
    - Application metrics
    - `MoistureSensor` application started
  - Receive *device commands* to alter behavior
    - Reduce maximum engine revs to reduce chance of failure
    - Icy Road Ahead!
    - Poll engine data every 200ms for five minutes
    - Start the `DeepEngineAnalysis` application

Side Note …

- There's no requirement to use this pattern and/or the device model

-

- Quarks could be used just to run local analytics for a self-contained control loop

-

- It's open source – if you find a new use for Quarks, great!
    - Join the community and promote it

Simple Scenario – Single Device View

Data Ingest

Analytics

IoTDevice
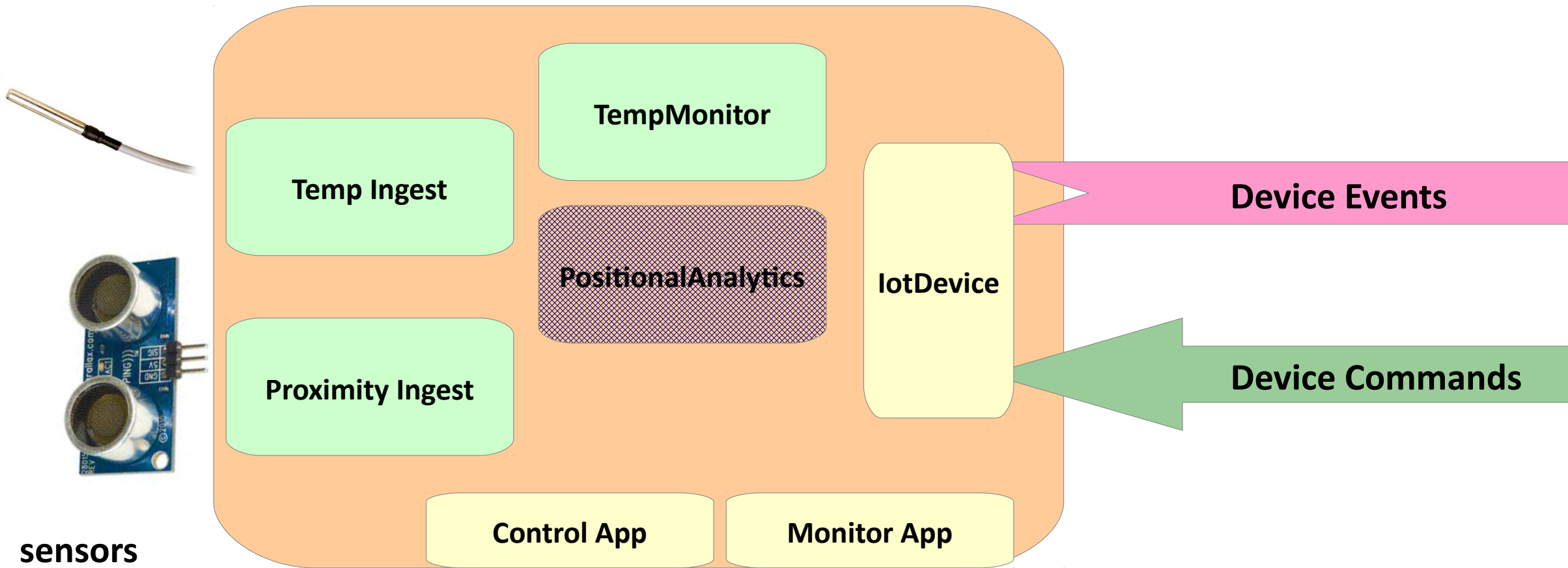
Device Events

Device Commands
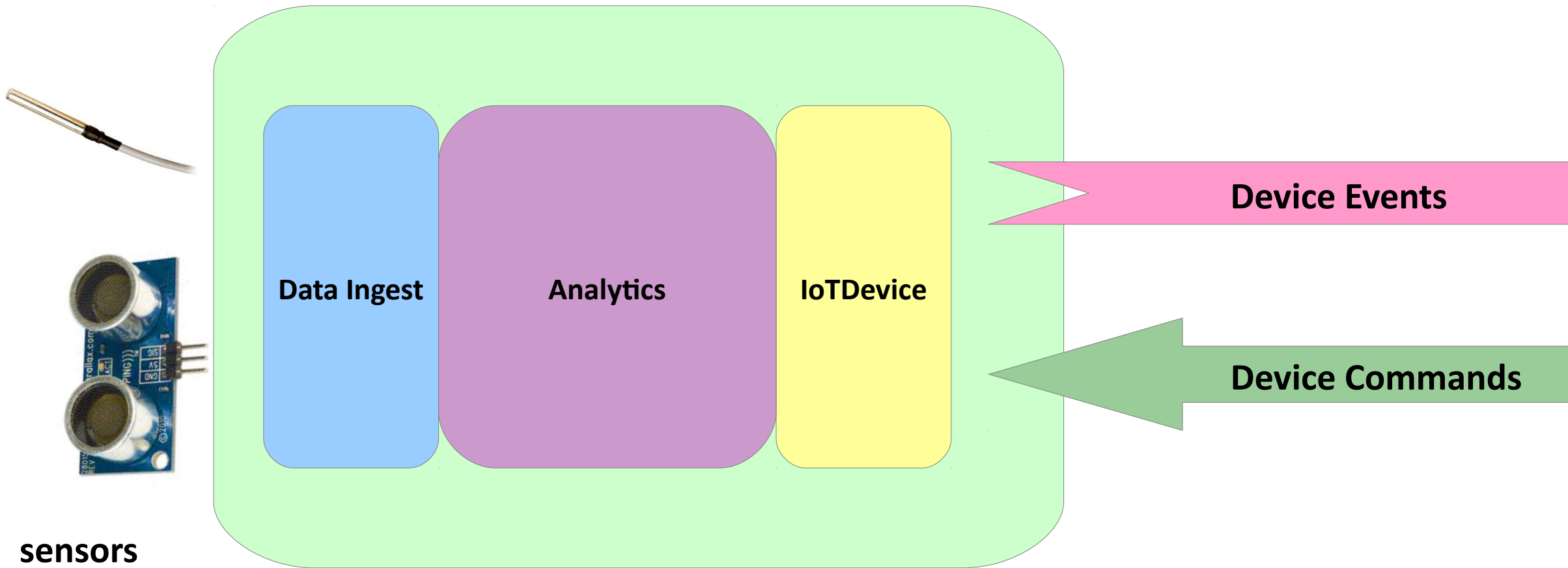
sensors

Quarks application (Topology)

Single Application

- Simplest use of Quarks

- Complete application in a single topology

- Can be limiting
  - Providers support multiple applications
  - MQTT based message hub may limit to single connection from authenticated device

- Good starting point
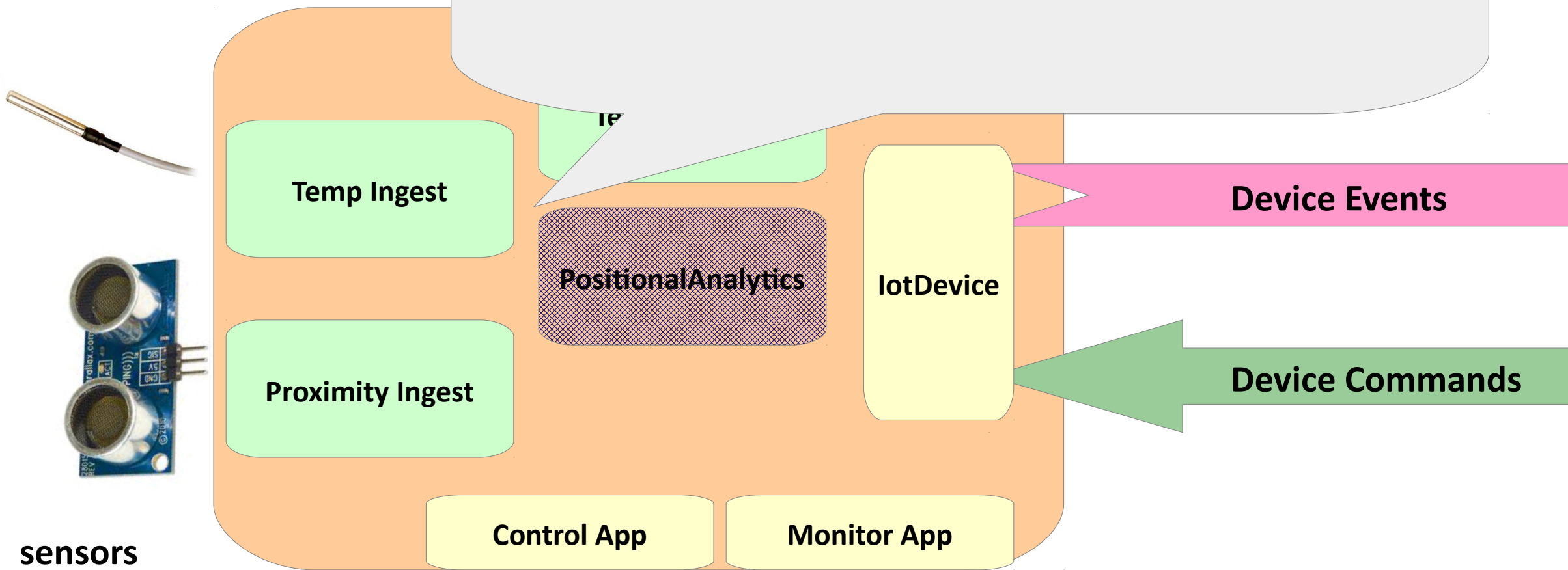 - Most samples are single applications

Full Scenario – Single Device View

sensors
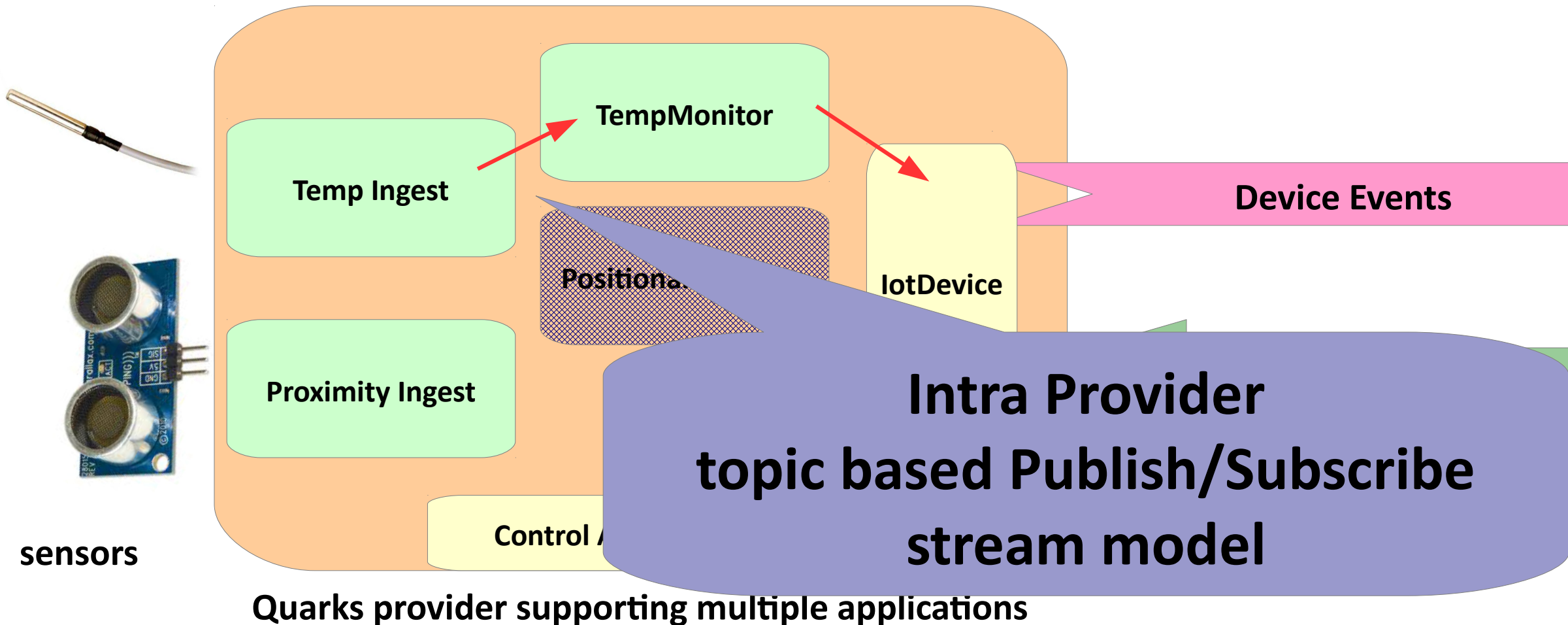
Quarks provider supporting multiple applications

TempMonitor

Temp Ingest

PositionalAnalytics

IotDevice

Proximity Ingest

Device Events

Device Commands

Control App

Monitor App

Full Scenar

Multiple independent applications

Temp Ingest

Proximity Ingest

PositionalAnalytics

IotDevice

Device Events

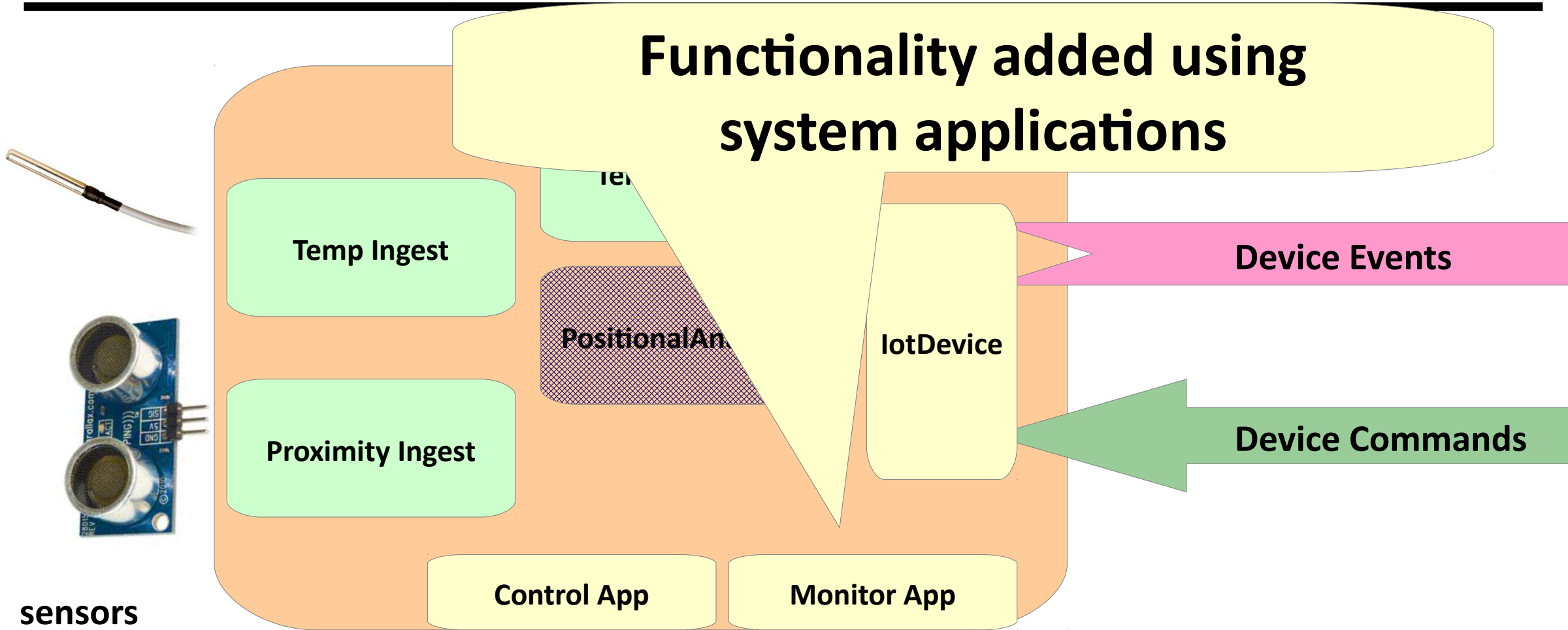**Device Commands**

Control App

Monitor App

sensors

Quarks provider supporting multiple applications

Publish Subscribe

TempMonitor

Temp Ingest

Positiona

IotDevice

Device Events

Proximity Ingest

**Intra Provider
topic based Publish/Subscribe
stream model**

Control A

sensors

Quarks provider supporting multiple applications

"System" Applications

Functionality added using system applications

Temp Ingest

Te...

PositionalAn...

IotDevice

Device Events

Proximity Ingest

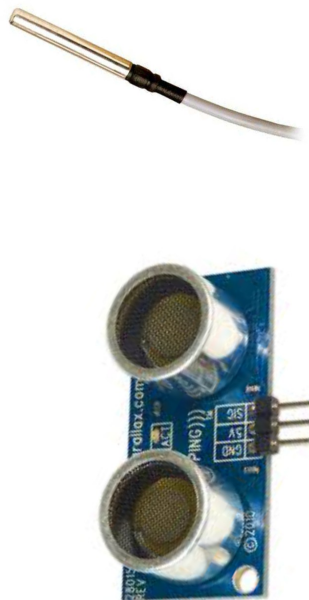**Device Commands**

Control App

Monitor App
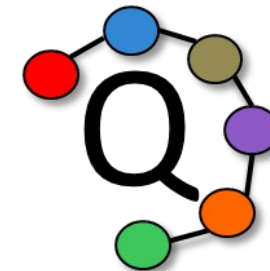
sensors

Quarks provider supporting multiple applications
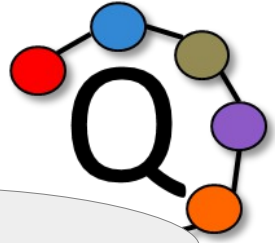
Application Registration
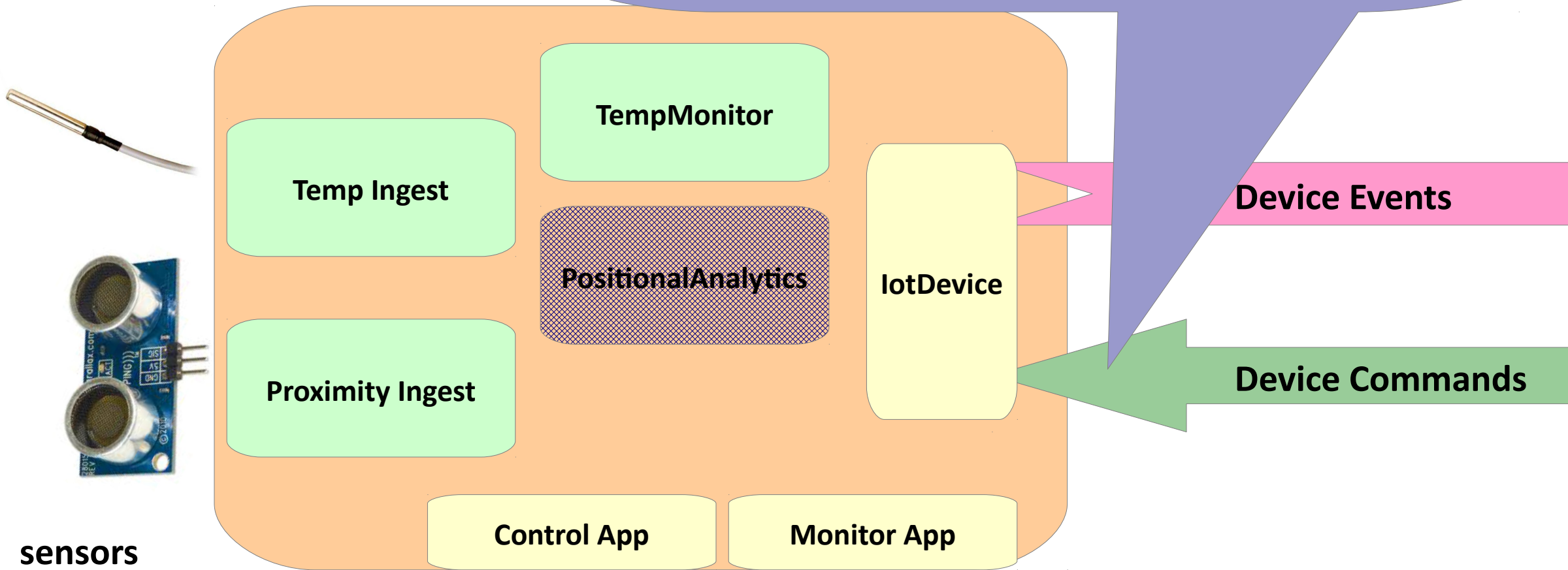
Applications can be registered without being started

Temp Ingest

Te...

PositionalAnalytics

IotDevice

Proximity Ingest

Control App

Monitor App

Device Events

Device Commands

sensors
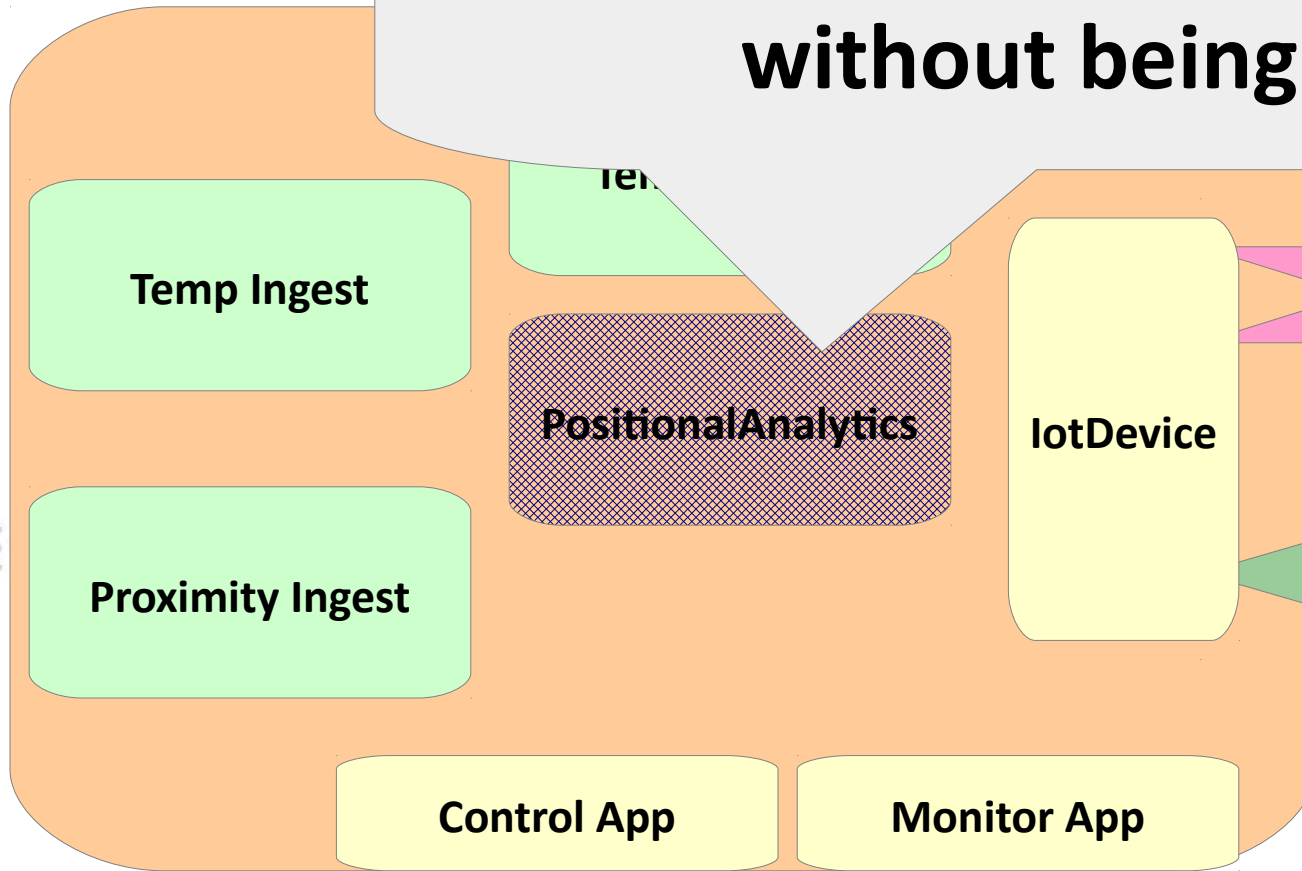
Quarks provider supporting multiple applications

Quarks Contro

Device commands can control applications

TempMonitor

Temp Ingest

PositionalAnalytics

Proximity Ingest

IotDevice

Device Events

Device Commands

Control App

Monitor App

sensors

Quarks provider supporting multiple applications

Start Application Device Command

Temp Ingest → **TempMonitor**

PositionalAnalytics

IotDevice → **Device Events**

Proximity Ingest

Control App   **Monitor App**

**Device Command
Start Application
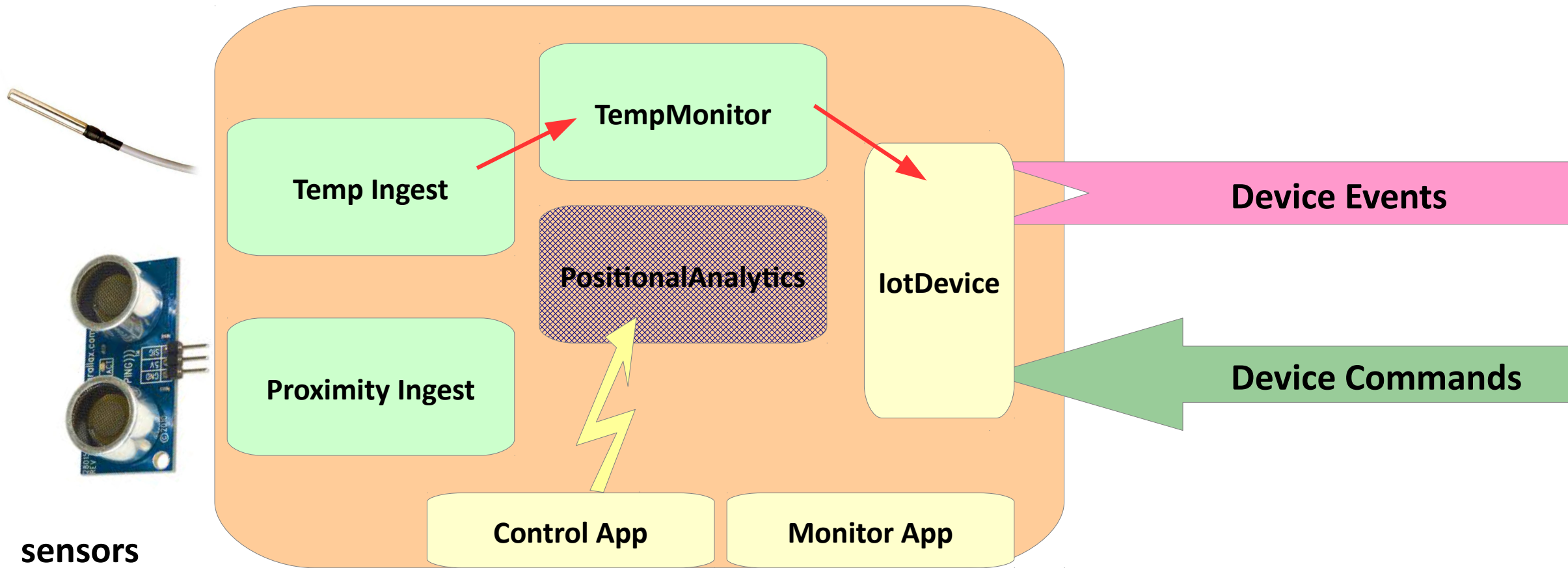PositionalAnalytics**

sensors
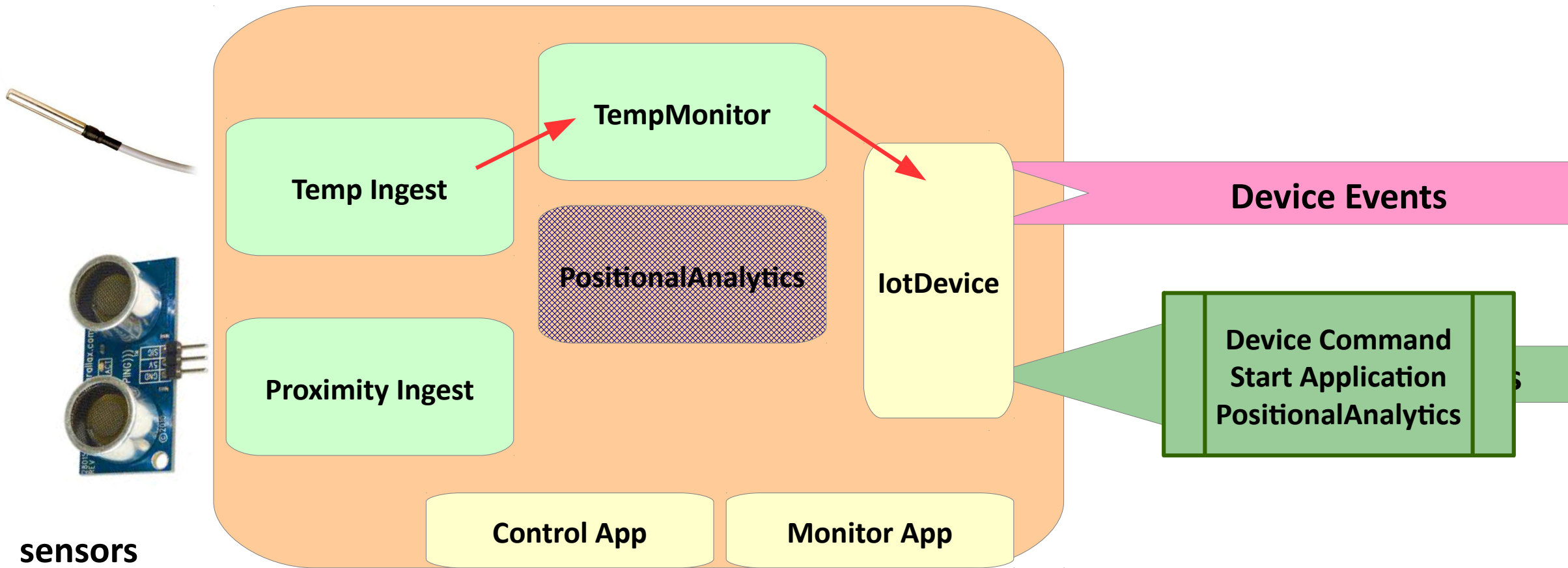
Quarks provider supporting multiple applications

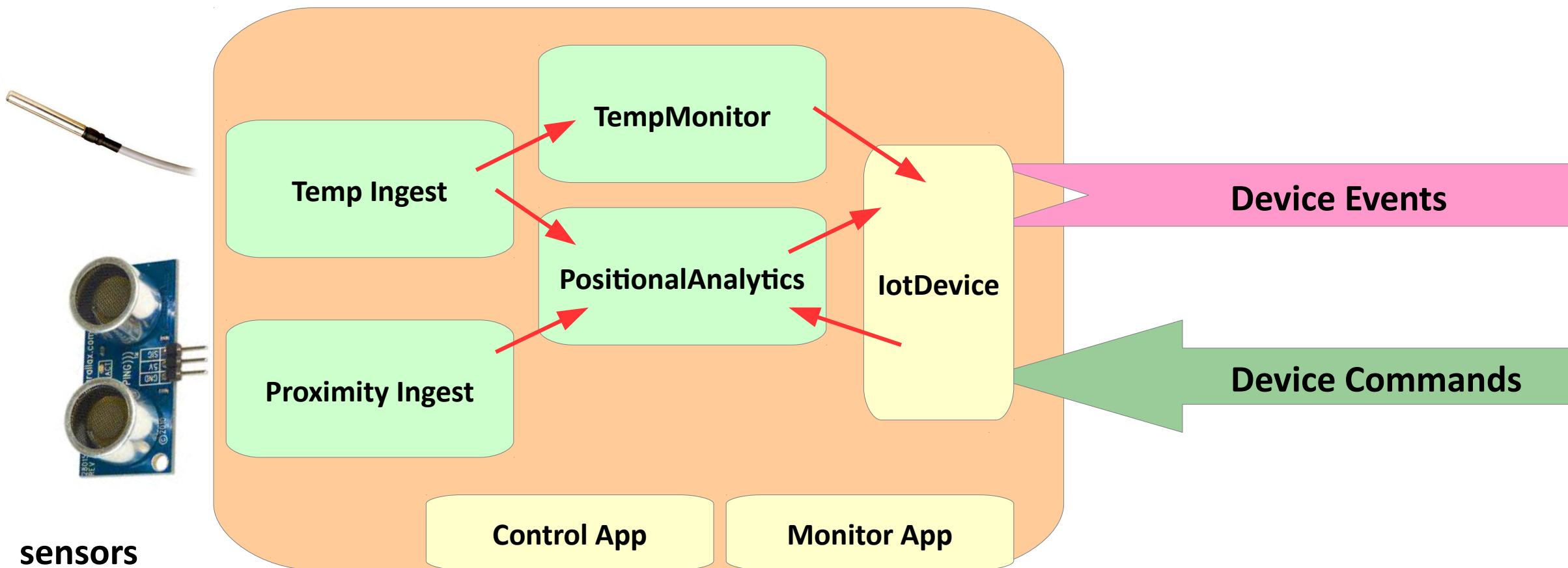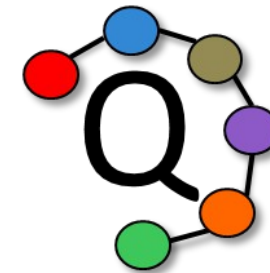# Control Application

sensors
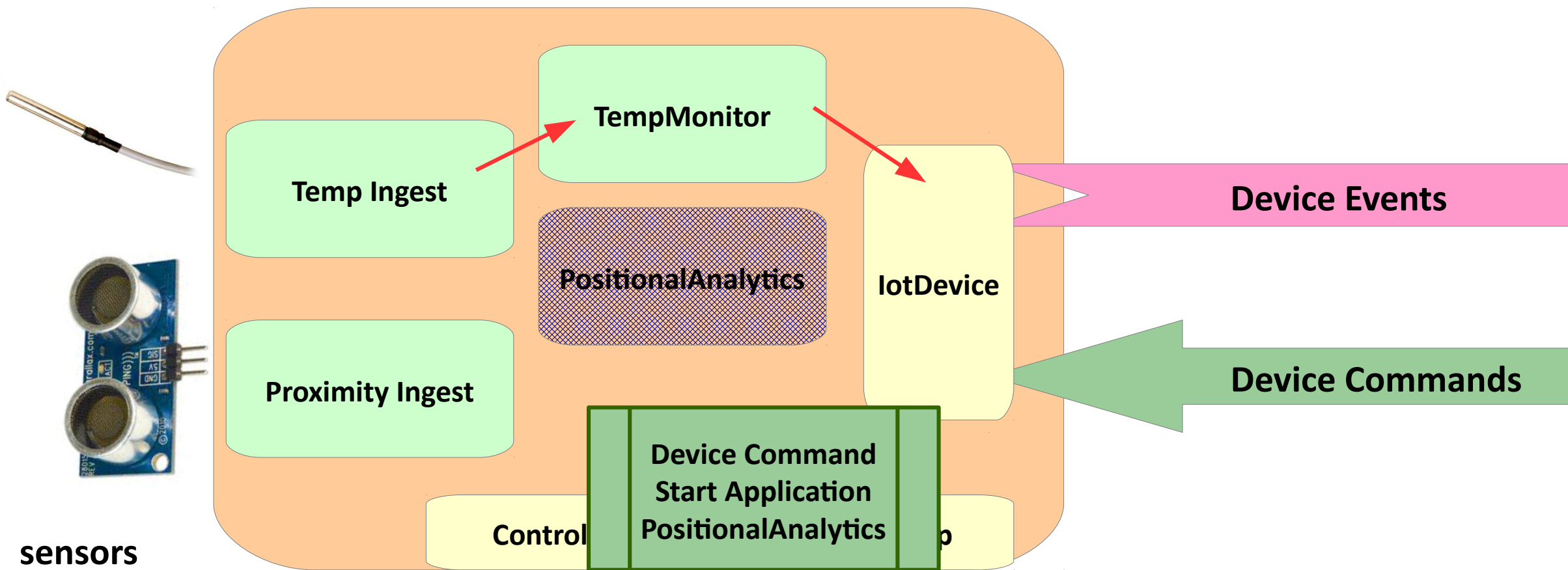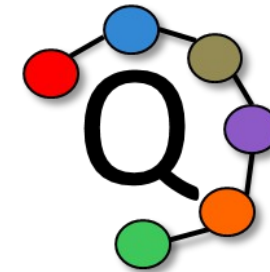
Quarks provider supporting multiple applications

Start Optional

sensors

Quarks provider supporting multiple applications

Application Running & Connected

sensors

Quarks provider supporting multiple applications

Full Scenario Use Cases

- Provide new service to clients
- Diagnose transient problem
  - Additional local analytics
  - Send additional data relevant to the problem
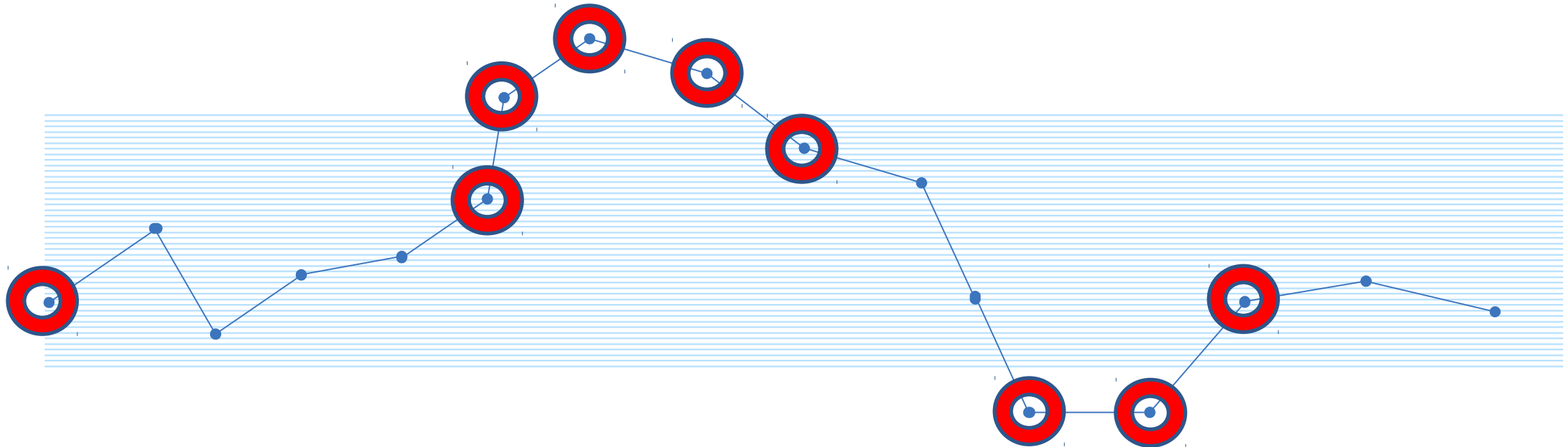- Temporary reduce resource consumption on constrained devices
- …

Quarks Features

- Functional Flow API for streaming analytics
  - Per-event and windowed processing with basic analytics
    - Map, FlatMap, Filter, Aggregates, Split, Union, Join, DeadbandFilter

- Connectors
  - Messaging systems & data stores
    - MQTT, HTTP, Web Sockets, JDBC, File, Apache Kafka and IBM Watson IoT Platform

- Micro-kernel style runtime with multi-platform support
  - Small-footprint edge devices or sensors
    - Including Raspberry Pis or smart phones

- Development mode
  - Web-console for viewing application graph and metrics

- Testing mechanism
  - Junit integration

Analytics

- Filtering, Map, FlatMap use functions (lambdas)
  - Can use any Java library that can run in targeted environment
  - `temps = temps.filter(t → t > 35.0)`

- Range

- Deadband filter

- Integration with Apache Common Math 3 provides:
  - Windowed aggregates, e.g. MAX, MIN, SUM, STDDEV, SLOPE
  - Kmeans+ clustering run on OBD-II demo on Raspberry Pi B
  - Investigation of exposing other algorithms for streaming

- Looking forward to additional contributions

Deadband Filter

Tuples within the band are discarded, except:
- – first return to being within band
- – more than T seconds since
  last tuple not discarded (optional)

```
temps = deadband(temps,
    j-> getDouble(j, "value"),
    v -> v > 5.0 && v < 25.0,
    20, SECONDS);
```

LAST SLIDE